Performance of Sphere Decoding of Block Codes

Mostafa El-Khamy, Haris Vikalo, Babak Hassibi, and Robert J. McEliece

***Abstract*—A sphere decoder searches for the closest lattice point within a certain search radius. The search radius provides a tradeoff between performance and complexity. We focus on analyzing the performance of sphere decoding of linear block codes. We analyze the performance of soft-decision sphere decoding on AWGN channels and a variety of modulation schemes. A hard-decision sphere decoder is a bounded distance decoder with the corresponding decoding radius. We analyze the performance of hard-decision sphere decoding on binary and q-ary symmetric channels. An upper bound on the performance of maximum-likelihood decoding of linear codes defined over F_q (e.g. Reed-Solomon codes) and transmitted over q-ary symmetric channels is derived and used in the analysis. We then discuss sphere decoding of general block codes or lattices with arbitrary modulation schemes. The tradeoff between the performance and complexity of a sphere decoder is then discussed.**

***Index Terms*—Maximum likelihood decoding, sphere decoding, performance bounds, Reed-Solomon codes, block codes, decoding radius, symmetric channels.**

I. Introduction

MAXIMUM likelihood (ML) decoding of linear block codes is known to be NP-hard [1]. A decoder that utilizes the soft output from the channel directly is called a *soft-decision* (SD) decoder. On the other hand, if hard decisions are made on the received bits before decoding, then such a decoder is called a *hard-decision* (HD) decoder. The optimum decoder is the corresponding HD or SD maximum likelihood (ML) decoder. Poltyrev derived tight upper bounds on the performance of maximum likelihood decoding of linear block codes over AWGN channels and binary symmetric (BSC) channels [2]. Berlekamp's tangential bound is a tighter bound than the union bound for additive white Gaussian noise (AWGN) channels [3]. Bounds based on typical pairs decoding were derived by Aji *et. al* [4]. Other bounds such as the Divsalar simple bound and the variations on the Gallager bounds are tight for AWGN and fading channels [5], [6]. For a broad survey on bounds on the performance of maximum likelihood decoding of linear codes, see [7].

Fincke and Pohst (FP) [8] described a sphere decoder algorithm which finds the closest lattice point without actually searching all the lattice points. A fast variation of it was given by Schnorr and Euchner [9]. Other efficient closest point search algorithms exist (for a survey see [10]). The sphere decoder algorithm was proposed for decoding lattice codes [11] and for detection in multiple antenna wireless systems [12], [13]. Vikalo and Hassibi proposed HD and SD sphere decoders for joint detection and decoding of linear block codes [14] [15]. On the other hand, one can think of a sphere decoder in a broader sense as any algorithm that returns the closest lattice point to the received word if it exists within a predetermined search radius. By this definition of a sphere decoder, the hard-decision Berlekamp-Massey algorithm can be considered as a sphere decoder for Reed Solomon (RS) codes with a search radius equal to half the minimum distance of the code. Similarly, the algorithm recently proposed by Guruswami and Sudan for decoding RS codes is an algebraic sphere decoder whose search radius can be larger than half the minimum distance of the code [16]. A sphere decoding algorithm, based on a reduced-state trellis decoding algorithm, was recently proposed in [17].

There has a been significant amount of research dedicated to the design of sphere decoders with smaller complexities, to the complexity analysis of sphere decoders and to the application of sphere decoders to various settings and communication systems. However, little research focused on the performance analysis of sphere decoders. This paper sets down a framework for the analysis of the performance of sphere decoding of block codes over a variety of channels with various modulation schemes.

In this paper, we study the performance of soft decision sphere decoding of linear block codes and lattices on channels with additive white Gaussian noise and various modulation schemes as BPSK, M-PSK and QAM [18]. This is done in sections II and III respectively. The application of these bounds to the binary image of Reed Solomon codes is also investigated. Bounds on the performance of hard decision sphere decoding on binary symmetric channels are derived in section IV. We then, in section V derive bounds on the maximum likelihood performance of q-ary linear codes, such as Reed Solomon codes, over q-ary symmetric channels. This bound becomes handy when analyzing the performance of sphere decoding of Reed Solomon codes on q-ary symmetric channels. Furthermore, we show, in section III, how one can analyze the performance of a soft decision sphere decoder of a general block code with a general modulation scheme. In many settings, we support our analytic bounds by comparing them to

Paper approved by K. Narayanan, the Editor for Coding and Communication Theory of the IEEE Communications Society. Manuscript received August 6, 2008.

M. El-Khamy is with the Electrical Engineering Department, Alexandria University, Alexandria, Egypt (e-mail: m_elkhamy@ieee.org).

H. Vikalo is with the Electrical and Computer Engineering Department, The University of Texas at Austin, Austin, TX (e-mail: hvikalo@ece.utexas.edu).

B. Hassibi and R. McEliece are with the Department of Electrical Engineering, California Institute of Technology, Pasadena, CA (e-mail: {hassibi, rjm}@systems.caltech.edu).

This research was supported by NSF grant no. CCF-0514881 and grants from Sony, Qualcomm, and the Lee Center for Advanced Networking. The work was performed while the authors were at the California Institute of Technology, Pasadena, CA. This work was presented in part at the 2005 IEEE Information Theory Workshop on Coding and Complexity, Rotorua, New Zealand and the 2006 IEEE International Symposium on Information Theory, Seattle, Washington.

Digital Object Identifier 10.1109/TCOMM.2009.10.080402

numerical simulations. The tradeoff between performance and complexity is discussed in section VI. Finally, we conclude our work in section VII.

II. Upper Bounds on the Performance of Soft Decision Sphere Decoding of BPSK and M-PSK Modulated Block Codes.

In this section, we consider a sphere decoder when the modulation is binary or M-ary phase shift keying (PSK) [18]. Each transmitted codeword in the code has the same energy when mapped to the PSK constellation. For the case of M-PSK modulation, complex sphere decoding algorithms which solve the closest point search problem were developed in [19].

A. Preliminaries

We will introduce some notation, so the bounds derived here are readily applicable for both M-ary and binary phase shift keying (PSK) modulation. We assume that $\mathcal{C}$ is an (n, k) linear code. Each codeword of length n will be mapped to a word of M PSK symbols. The number of channel symbols will be denoted by n_c. If the code $\mathcal{C}$ is binary and of length n, then $n_c = \lceil n/\log_2(M) \rceil$. For BPSK, $n_c = n$. Note that the original code need not be binary. For example, an Reed Solomon (RS) code defined over $\mathbb{F}_{2^m}$ could be mapped directly to an 2^m-ary PSK constellation by a one-to-one mapping from the symbols in $\mathbb{F}_{2^m}$ to the 2^m points in the PSK constellation.

For PSK signaling, the code will have the property that all codewords are of equal energy and lie on a sphere of radius $\sqrt{n_c}$ from the origin of space. Let n_d denote the dimension of the considered space (noise). For the case of BPSK modulation, the dimension of the Hamming space is the same as the number of channel symbols (bits) $n_d = n_c$. On the other hand, for M-PSK signaling, $M > 2$, each complex channel symbol has a real and an imaginary component. Thus the noise has $2\,n_c$ independent components and the dimension of the space is $n_d = 2\,n_c$.

Assuming that a codeword $\boldsymbol{c} \in \mathcal{C}$ is transmitted over a binary input AWGN channel, the received word is $\boldsymbol{y} = \boldsymbol{x} + \boldsymbol{z}$, where $\boldsymbol{x} = \mathcal{M}(\boldsymbol{c})$ and $\mathcal{M}(\boldsymbol{c})$ is the mapping of the codeword $\boldsymbol{c}$ under PSK modulation, i.e., for BPSK modulation $\mathcal{M}(\boldsymbol{c}) \stackrel{\Delta}{=} 1 - 2\boldsymbol{c}$. The additive white Gaussian noise (AWGN) is denoted by $\boldsymbol{z} = [z_i]_{i=1}^{n_d}$ with variance σ^2. Let G_w be the number of codewords which (after mapping) are at an Euclidian distance δ_w from each other. Note that for the case of BPSK modulation and a binary code $\mathcal{C}$, the space is a Hamming space and the Euclidean distance is directly related to the Hamming distance, $\delta_w = 2\sqrt{w}$, where w is the Hamming distance. QPSK modulation with Gray encoding also results in a Hamming space [18] by $\delta_w = \sqrt{2w}$, where w is the (binary) Hamming distance between the codewords. In the following analysis, it is assumed that the modulated code is linear and the transmitted signal set is assumed to be geometrically uniform [20] where the conditional error probability does not depend on the transmitted signal point (codeword).

B. Analysis of Soft Decision Sphere Decoding

A soft-decision sphere decoder with an Euclidean radius D, denoted by SSD(D), solves the following optimization problem,

$$\begin{aligned} \hat{\boldsymbol{c}} = \arg\min_{\boldsymbol{v}\in\mathcal{C}} \quad & \|\boldsymbol{y} - \mathcal{M}(\boldsymbol{v})\|^2 \\ \text{subject to} \quad & \|\boldsymbol{y} - \mathcal{M}(\boldsymbol{v})\|^2 \le D^2, \end{aligned} \tag{1}$$

where $\|\boldsymbol{x}\|$ is the Euclidean norm of $\boldsymbol{x}$. Such decoders include *list-decoders* that list all codewords whose modulated image is within an Euclidean distance D from the received vector y and choose the closest one. If no such codeword exists, a decoding *failure* is signaled. A decoding *error* is signaled if the decoded codeword is not the transmitted codeword.

Let $\mathcal{E}_D$ denote the event of error or failure of SSD(D), then the error plus failure probability, $P(\mathcal{E}_D)$ [1] is

$$P(\mathcal{E}_D) = P(\mathcal{E}_D|\mathcal{E}_{ML})P(\mathcal{E}_{ML}) + P(\mathcal{E}_D|\mathcal{S}_{ML})P(\mathcal{S}_{ML}), \tag{2}$$

where $\mathcal{E}_{ML}$ and $\mathcal{S}_{ML}$ denote the events of an ML error and an ML success respectively. Let $\epsilon = \|\boldsymbol{y} - \mathcal{M}(\boldsymbol{c})\|$, then an ML error results if there exists another codeword $\hat{\boldsymbol{c}} \in \mathcal{C}$ such that $\|\boldsymbol{y} - \mathcal{M}(\hat{\boldsymbol{c}})\| \le \epsilon$. Since limiting the decoding radius to D will not do better than ML decoding, then $P(\mathcal{E}_D|\mathcal{E}_{ML}) = 1$. By observing that $P(\mathcal{S}_{ML}) \le 1$, it follows that an upper bound on the decoding performance is

$$P(\mathcal{E}_D) \le P(\mathcal{E}_{ML}) + P(\mathcal{E}_D|\mathcal{S}_{ML}). \tag{3}$$

Let Ω_D be the Euclidean sphere of radius D centered around the transmitted codeword in the n_d dimensional space. The probability that the added white Gaussian noise will not lie in the sphere Ω_D is

$$\begin{aligned} P(\boldsymbol{z} \notin \Omega_D) &= P\left(\chi_{n_d} > D^2\right) \\ &= 1 - \Gamma_r(n_d/2, D^2/2\sigma^2) \end{aligned} \tag{4}$$

where $\chi_n = \sum_{i=1}^{n} z_i^2$ is a Chi-squared distributed random variable with n degrees of freedom. Let $\Gamma(x)$ denote the Gamma function, then the cumulative distribution function (CDF) of χ_v is given by the regularized Gamma function Γ_r [21],

$$\Gamma_r(v/2, w/2) = \begin{cases} \int_0^w \frac{t^{v/2-1}e^{-t/2}}{2^{v/2}\Gamma(v/2)}\mathrm{d}t, & w \ge 0 \\ 0, & w < 0. \end{cases} \tag{5}$$

Lemma 1: A lower bound on $P(\mathcal{E}_D)$ is $P(\mathcal{E}_D) \ge P(\boldsymbol{z} \notin \Omega_D)$.

Proof: The sphere decoder error plus failure probability could be written as

$$\begin{aligned} P(\mathcal{E}_D) &= P(\mathcal{E}_D|\boldsymbol{z} \in \Omega_D)P(\boldsymbol{z} \in \Omega_D) \\ &\quad + P(\mathcal{E}_D|\boldsymbol{z} \notin \Omega_D)P(\boldsymbol{z} \notin \Omega_D) \\ &\ge P(\mathcal{E}_D|\boldsymbol{z} \notin \Omega_D)P(\boldsymbol{z} \notin \Omega_D) \\ &= P(\boldsymbol{z} \notin \Omega_D), \end{aligned}$$

where the last inequality is because $P(\mathcal{E}_D|\boldsymbol{z} \notin \Omega_D) = 1$ which follows from the definition of the sphere decoder (1). ■

[1] Through out this paper, $P(X)$ will denote the probability that the event X occurs.

Define $\bar{P}(\mathcal{E}_{ML})$ to be an upper bound on the SD-ML decoder error probability, then we have the following lemma,

Lemma 2: $P(\mathcal{E}_D) \leq \bar{P}(\mathcal{E}_{ML}) + P(\boldsymbol{z} \notin \Omega_D)$.

Proof: Given an ML success, $\mathcal{E}_D$ will only be due to failures of the SSD(D) decoder, i.e.,

$$P(\mathcal{E}_D|\mathcal{S}_{ML}) = P(\|\boldsymbol{y} - \mathcal{M}(\boldsymbol{c})\| > D) = P(\boldsymbol{z} \notin \Omega_D),$$

because the channel is additive noise and memoryless. By definition, $P(\mathcal{E}_{ML}) \leq \bar{P}(\mathcal{E}_{ML})$. By substituting in (3) we are done. ■

Lemma 2 provides a way to bound the performance of sphere decoding of linear block codes on a variety of channels where additive white Gaussian noise is added and for a variety of modulation schemes.

Lemma 1 implies that one could obtain a tighter upper bound on $P(\mathcal{E}_D)$ by tightening the bound on the ML error probability, $\bar{P}(\mathcal{E}_{ML})$. Shannon's sphere packing bound [22] is a lower bound on the error probability where Shannon showed that the Voronoi region of a codeword can be bounded by a right circular n_d-dimensional cone with the codeword on its axis. Poltyrev's tangential sphere bound (TSB) is one of the tightest bounds on the ML performance of soft decision decoding of linear codes on AWGN channels with BPSK or M-PSK modulation [2], [23] and is calculated by,

$$P(\mathcal{E}_{ML}) \leq \min_{\theta} \left\{ P(\mathcal{E}_{ML}, \boldsymbol{z} \in V_\theta) + P(\boldsymbol{z} \notin V_\theta) \right\}, \qquad (6)$$

where V_θ is an n_d-dimensional right circular cone with a half angle θ whose central line passes through the transmitted codeword and whose apex is at an Euclidean distance $\sqrt{n_c}$ from the transmitted codeword. Let the minimum of the optimization problem in (6) be achieved at $\theta = \phi$, then by Lemma 2 we have the following upper bound

$$P(\mathcal{E}_D) \leq P(\mathcal{E}_{ML}, \boldsymbol{z} \in V_\phi) + P(\boldsymbol{z} \notin V_\phi) + P(\boldsymbol{z} \notin \Omega_D). \qquad (7)$$

For the TSB, the optimum angle ϕ is related to the radius $\sqrt{r_\phi}$ (see Fig. 1 or Fig. 2) by $\tan(\phi) = \sqrt{r_\phi / n_c}$, such that r_ϕ is the root of this equation [23]

$$\sum_{\delta_b > 0} G_b'(r_o) \int_0^{\theta_b(r_o)} \sin^{n_d - 3}(\vartheta) \mathrm{d}\vartheta = \frac{\sqrt{\pi}\Gamma(\frac{n_d-2}{2})}{\Gamma(\frac{n_d-1}{2})} \qquad (8)$$

when solved for r_o, where $\theta_b(r_o) \triangleq \cos^{-1}\left(\frac{\delta_b/2}{\sqrt{r_o(1-\delta_b^2/4n_c)}}\right)$, and

$$G_b'(r_o) = \begin{cases} G_b, & \delta_b^2/4 < r_o(1 - \delta_b^2/4n_c) \\ 0, & \text{otherwise.} \end{cases} \qquad (9)$$

Let z_1 be the component of the noise along the central axis of the cone with a probability distribution function (PDF) $\mathcal{N}(z_1) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-{z_1}^2/2\sigma^2}$ and z_2 be the noise component orthogonal to z_1. Define $\beta_{z_1}(w) \triangleq \frac{\sqrt{n_c} - z_1}{\sqrt{\frac{4n_c}{\delta_w^2} - 1}}$ and $r_{z_1}(\phi) \triangleq \sqrt{r_\phi}\left(1 - \frac{z_1}{\sqrt{n_c}}\right)$, then the ML error probability given that the noise $\boldsymbol{z}$ is in the cone V_ϕ is [2]

$$P(\mathcal{E}_{ML}, \boldsymbol{z} \in V_\phi) = \int_{-\infty}^{\infty} \mathcal{N}(z_1) \left[\sum_{\delta_b > 0} G_b'(r_\phi) \int_{\beta_{z_1}(b)}^{r_{z_1}(\phi)} \mathcal{N}(z_2) \Gamma_r\left(\frac{n_d-2}{2}, \frac{r_{z_1}^2(\phi) - z_2^2}{2\sigma^2}\right) \mathrm{d}z_2 \right] \mathrm{d}z_1. \qquad (10)$$

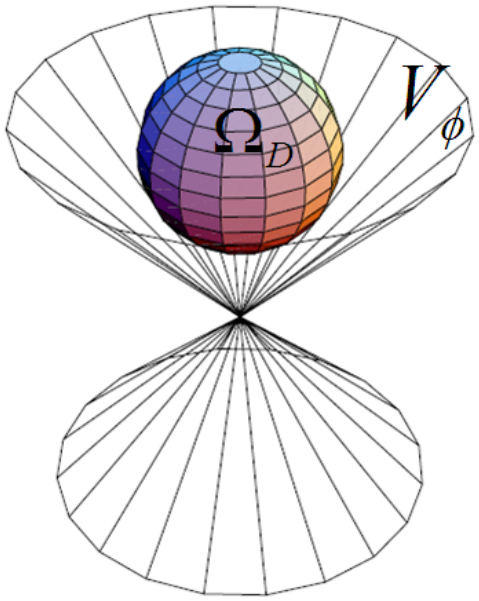

Fig. 1. *Case A:* The sphere Ω_D lies totally inside the Cone V_ϕ ($D \leq \sqrt{n_c}\sin(\phi)$).

C. A Tight Upper Bound

We observe that instead of directly substituting the TSB of (6) for $\bar{P}(\mathcal{E}_{ML})$ in Lemma 2 as we did in (7), one can find an upper bound which is tighter than (7) by noticing that the events $\{\boldsymbol{z} \notin V_\theta\}$ and $\{\boldsymbol{z} \notin \Omega_D\}$ are not in general mutually exclusive.

Lemma 3: $P(\mathcal{E}_D)$ is upper bounded by

$$P(\mathcal{E}_D) \leq P(\mathcal{E}_{ML}, \boldsymbol{z} \in V_\phi) + P(\boldsymbol{z} \notin \Omega_D) + P\left(\{\boldsymbol{z} \notin V_\phi\} \cap \{\boldsymbol{z} \in \Omega_D\}\right).$$

Proof: Using Bayes' rule and defining the region $\Lambda(\theta, D) \triangleq \{V_\theta \cap \Omega_D\}$ we get

$$P(\mathcal{E}_D) \leq \min_{\theta} \{ P(\mathcal{E}_D|\boldsymbol{z} \in \Lambda(\theta, D)) P(\boldsymbol{z} \in \Lambda(\theta, D)) + P(\mathcal{E}_D|\boldsymbol{z} \notin \Lambda(\theta, D)) P(\boldsymbol{z} \notin \Lambda(\theta, D)) \}. \qquad (11)$$

From the definition of $\Lambda(\theta, D)$, it follows that $P(\mathcal{E}_D, \boldsymbol{z} \in \Lambda(\theta, D)) = P(\mathcal{E}_{ML}, \boldsymbol{z} \in \Lambda(\theta, D)) \leq P(\mathcal{E}_{ML}, \boldsymbol{z} \in V_\theta)$, where the last inequality follows from that $\Lambda(\theta, D) \subseteq V_\theta$. Using $P(\mathcal{E}_D|\boldsymbol{z} \notin \Lambda(\theta, D)) \leq 1$, it follows that

$$P(\mathcal{E}_D) \leq \min_{\theta} \{ P(\mathcal{E}_{ML}, \boldsymbol{z} \in V_\theta) + P(\boldsymbol{z} \notin \Lambda(\theta, D)) \} \leq P(\mathcal{E}_{ML}, \boldsymbol{z} \in V_\phi) + P(\boldsymbol{z} \notin \{V_\phi \cap \Omega_D\}). \qquad (12)$$

The last inequality is due to the observation that ϕ does not necessarily minimize (12). By de Morgan's law, $\{V_\phi \cap \Omega_D\}^c = \{\Omega_D\}^c \cup \{\{V_\phi\}^c \cap \Omega_D\}$, $\{.\}^c$ is the complement of $\{.\}$. ■

We consider two cases;

Case A: The sphere Ω_D lies totally inside the cone V_ϕ. (See Fig. 1). This case is equivalent to the event $\mathbb{A} \triangleq \{D \leq D_\phi\}$, where

$$D_\phi = \sqrt{n_c}\sin(\phi), \qquad (13)$$

and will be called the critical decoding radius. It follows that $P\left(\{\boldsymbol{z} \notin V_\phi\} \cap \{\boldsymbol{z} \in \Omega_D\}|\mathbb{A}\right) = 0$, which could be substituted

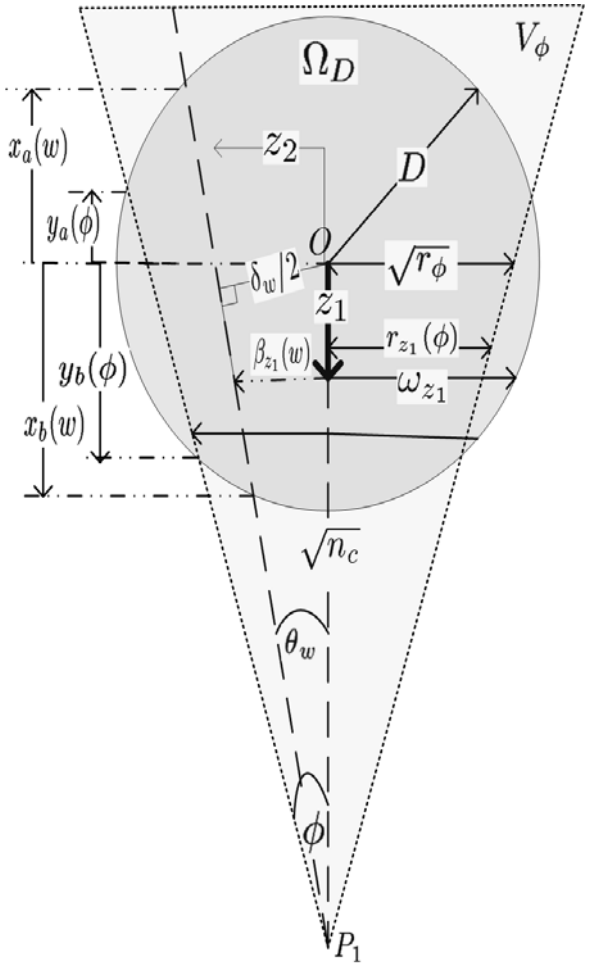

Fig. 2. *Case B:* The sphere Ω_D intersects the cone V_ϕ: the apex of the cone V_ϕ lies outside the sphere Ω_D if $\sqrt{n_c}\sin(\phi) < D < \sqrt{n_c}$, the apex of the cone V_ϕ lies inside the sphere Ω_D if $D \geq \sqrt{n_c}$.

in Lemma 2. Furthermore, since $\Lambda(\theta, D) = \Omega_D$, it follows from (11) that a tighter upper bound is

$$P(\mathcal{E}_D|\mathbb{A}) \leq P(\mathcal{E}_{ML}, \boldsymbol{z} \in \Omega_D) + P(\boldsymbol{z} \notin \Omega_D). \tag{14}$$

The joint probability of the added noise falling inside a sphere of Euclidean radius D and an ML error could be expressed as

$$P(\mathcal{E}_{ML}, \boldsymbol{z} \in \Omega_D) = \sum_{0<\frac{\delta_b}{2}<D} G_b \int_{\frac{\delta_b}{2}}^{D} \mathcal{N}(z_o)\Gamma_r\left(\tfrac{n_d-1}{2}, \tfrac{D^2-z_o^2}{2\sigma^2}\right) \mathrm{d}z_o. \tag{15}$$

Let φ be the half angle at which the cone V_φ is tangential to the sphere Ω_D, $\varphi = \sin^{-1}(D/\sqrt{n})$ (see Fig. 1), then another tight upper bound is

$$P(\mathcal{E}_D|\mathbb{A}) \leq P(\mathcal{E}_{ML}, \boldsymbol{z} \in V_\varphi) + P(\boldsymbol{z} \notin \Omega_D). \tag{16}$$

Theoretically, it is clear that the bound of (14) is tighter than that of (16), but numerically they are almost equivalent, since the integration over the region $\{\Omega_D^c \bigcap V_\varphi\}$ is negligible. Note that $P(\mathcal{E}_{ML}, \boldsymbol{z} \in V_\varphi)$ is easily calculated using equation (10) where $\tan(\varphi) = \sqrt{r_\varphi/n_c}$ and $r_{z_1}(\varphi) = \sqrt{r_\varphi}\left(1 - \frac{z_1}{\sqrt{n_c}}\right)$. □

Case B: The sphere Ω_D intersects the cone V_ϕ. (see Fig. 2). We have two cases depending on the position of the apex of the cone. The first is when the apex of the cone does not lie in the sphere, $\sqrt{n_c}\sin(\phi) < D < \sqrt{n_c}$ and the second is when the apex lies in the sphere, $D \geq \sqrt{n_c}$ (see Fig. 2). In both cases the following analysis holds. Let the origin, O, of the n_d-dimensional space be at the transmitted codeword which is also the center of Ω_D. Since the cone and the sphere are symmetrical around the central axis, we project on a two dimensional plane as in Fig. 2. The radial component of the noise (along the axis of the cone) is z_1. The altitudes $y_a(\phi)$ and $y_b(\phi)$ at which the (double) cone intersects the sphere are found by substituting the line equation $P = P_1 + U(P_2 - P1)$, where $P = (x, y)$, $P_1 = (0, \sqrt{n_c})$ and $P_2 = (2\sqrt{n_c}\tan(\phi), -\sqrt{n_c})$ into the quadratic equation of the sphere. It follows that $y_{a,b}(\phi) = \sqrt{n_c}(1 - 2U_{a,b}(\phi, D))$, where

$$U_{a,b}(\theta, D) = \frac{4n_c \pm \sqrt{16{n_c}^2 - 16n_c\sec^2(\theta)(n_c - D^2)}}{8n_c\sec^2(\theta)}.$$

It is easy to check that at $D = \sqrt{n_c}$, $u_b = 0$ and y_b is at the apex of V_ϕ. If $D > \sqrt{n_c}$ then the intersection at y_b is in the lower nappe of the cone. It is also observed that V_ϕ and Ω_D do not intersect ($\Omega_D \subset V_\phi$) if $16n_c^2 < 16n_c\sec^2(\phi)(n_c - D^2)$ or equivalently $D < \sqrt{n_c}\sin(\phi)$ which is Case A.

Define $\mathbb{B}$ to be the event $\mathbb{B} \triangleq \left\{D > \sqrt{n_c}\sin(\phi)\right\}$, $f_{n-1}(t)$ to be the PDF of $\chi_{n-1} = \sum_{i=2}^{n} z_i^2$, and $\omega_{z_1}^2 = D^2 - z_1^2$ (see Fig. 2). From Lemma 3, the error probability is upper bounded by

$$\begin{aligned} P(\mathcal{E}_D|\mathbb{B}) &\leq P(\mathcal{E}_{ML}, \boldsymbol{z} \in V_\phi) \\ &\quad + P(\boldsymbol{z} \notin \Omega_D) + P\left(\{\boldsymbol{z} \notin V_\phi\} \cap \{\boldsymbol{z} \in \Omega_D\}|\mathbb{B}\right), \end{aligned} \tag{17}$$

where by Fig. 2

$$\begin{aligned} &P\left(\{\boldsymbol{z} \notin V_\phi\} \cap \{\boldsymbol{z} \in \Omega_D\}|\mathbb{B}\right) = \\ &\quad \int_{y_a(\phi)}^{y_b(\phi)} \mathcal{N}(z_1) \int_{r_{z_1}^2(\phi)}^{\omega_{z_1}^2} f_{n_d-1}(t)\mathrm{d}t\mathrm{d}z_1. \end{aligned} \tag{18}$$

□

The tight upper bound is summarized in this theorem,

Theorem 4: The performance of soft decision sphere decoding with an Euclidean decoding radius D of a linear code with (Euclidean) weight spectrum G_b on an AWGN channel with noise variance σ^2 and (binary or M-ary) PSK modulation is upper bounded by:

$$P(\mathcal{E}_D) \leq \begin{cases} \sum_{0<\frac{\delta_b}{2}<D} G_b \int_{\frac{\delta_b}{2}}^{D} \frac{e^{-z_o^2/2\sigma^2}}{\sqrt{2\pi\sigma^2}} \Gamma_r\left(\tfrac{n_d-1}{2}, \tfrac{D^2-z_o^2}{2\sigma^2}\right) \mathrm{d}z_o \\ +1 - \Gamma_r(n_d/2, D^2/2\sigma^2), \\ \text{if } D \leq \sqrt{n_c}\sin(\phi), \\ \\ \int_{-\infty}^{\infty} \mathcal{N}(z_1) \left[\sum_{\delta_b>0} G_b'(r_\phi) \int_{\beta_{z_1}(b)}^{r_{z_1}(\phi)} \mathcal{N}(z_2) \right. \\ \left. \Gamma_r\left(\tfrac{n_d-2}{2}, \tfrac{r_{z_1}^2(\phi)-z_2^2}{2\sigma^2}\right) \mathrm{d}z_2\right] \mathrm{d}z_1 + 1 \\ -\Gamma_r(n_d/2, D^2/2\sigma^2) + \int_{y_a(\phi)}^{y_b(\phi)} \left(\Gamma_r\left(\tfrac{n_d-1}{2}, \tfrac{\omega_{z_1}^2}{2\sigma^2}\right)\right. \\ \left. -\Gamma_r\left(\tfrac{n_d-1}{2}, \tfrac{r_{z_1}^2(\phi)}{2\sigma^2}\right)\right) \mathcal{N}(z_1)\mathrm{d}z_1, \\ \text{if } D > \sqrt{n_c}\sin(\phi), \end{cases}$$

where ϕ is the half angle of the cone V_ϕ and is given by (8). ▽

Following the proof of Lemma 3, the error plus failure probability of SSD(D) is upper bounded by

$$P(\mathcal{E}_D) \le P(\mathcal{E}_D, \boldsymbol{z} \in \Lambda(\phi, D)) + P(\boldsymbol{z} \notin \Lambda(\phi, D)). \quad (19)$$

From the previous arguments in *Cases A* and *B*, the following theorem provides a slightly tighter upper bound than that of the previous theorem.

Theorem 5: The performance of SSD(D) for BPSK or MPSK modulation is upper bounded by

$$P(\mathcal{E}_D) \le \begin{cases} P(\mathcal{E}_{ML}, \boldsymbol{z} \in \Omega_D) + P(\boldsymbol{z} \notin \Omega_D), \\ \text{if } D \le \sqrt{n_c}\sin(\phi); \\ \\ P(\mathcal{E}_{ML}, \boldsymbol{z} \in \Lambda(\phi, D)) + P(\boldsymbol{z} \notin \Omega_D) + \\ P\left(\{\boldsymbol{z} \notin V_\phi\} \cap \{\boldsymbol{z} \in \Omega_D\}\right), \\ \text{if } D > \sqrt{n_c}\sin(\phi) \end{cases}.$$

$\triangledown$

Observe that the difference from Theorem 4 is that the term $P(\mathcal{E}_{ML}, \boldsymbol{z} \in \Lambda(\phi, D))$ was upper bounded by $P(\mathcal{E}_{ML}, \boldsymbol{z} \in V(\phi))$ in Theorem 4. Consider a codeword at a distance δ_w, then the half angle of the cone bisecting this distance is $\theta_w = \sin^{-1}(\delta_w/2\sqrt{n_c})$ (c.f. Fig. 2). This cone will intersect the sphere Ω_D at altitudes $x_a(w)$ and $x_b(w)$ given by $x_{a,b}(w) = \sqrt{n_c}(1 - 2\,U_{a,b}(\theta_w, D))$. Now define the integrals

$$\mathcal{I}(\gamma, w, z_1) \triangleq \mathcal{N}(z_1) \int_{\beta_{z_1}(w)}^{\gamma} \mathcal{N}(z_2)\Gamma_r\left(\tfrac{n_d-2}{2}, \tfrac{\gamma^2 - z_2^2}{2\sigma^2}\right) \mathrm{d}z_2 \quad (20)$$

$$\mathcal{I}_2(w) = \int_{x_a(w)}^{y_a(\phi)} \mathcal{I}(\omega_{z_1}, w, z_1)\mathrm{d}z_1 + \int_{y_a(\phi)}^{y_b(\phi)} \mathcal{I}(r_{z_1}(\phi), w, z_1)\mathrm{d}z_1 + \int_{y_b(\phi)}^{x_b(w)} \mathcal{I}(\omega_{z_1}, w, z_1)\mathrm{d}z_1. \quad (21)$$

Taking the union over all codewords with non-zero Euclidean weights such that $\theta_w < \phi$, it follows that for $D > \sqrt{n_c}\sin(\phi)$,

$$P(\mathcal{E}_{ML}, \boldsymbol{z} \in \Lambda(\phi, D)) = \sum_{\delta_b > 0} G_b'(r_\phi)\mathcal{I}_2(w). \quad (22)$$

Observation: The bound of Theorem 5 is upper bounded by 1. This bound can be written as $P(\mathcal{E}_D) \le P(\mathcal{E}_{ML}, \boldsymbol{z} \in \Lambda(\phi, D)) + P(\boldsymbol{z} \notin \Lambda(\phi, D)) \le P(\boldsymbol{z} \in \Lambda(\phi, D)) + P(\boldsymbol{z} \notin \Lambda(\phi, D)) = 1$.

D. A Note on Reed-Solomon Codes

Consider the case when the binary image of an Reed-Solomon (RS) code, defined over $\mathbb{F}_{2^m}$, is transmitted over an AWGN channel and the decoder is either a HD or SD sphere decoder. The weight enumerator of an ensemble of binary images of generalized RS codes was derived by Retter [24]. Tight upper bounds on the performance of HD and SD maximum likelihood decoding of the binary images of RS codes were developed by El-Khamy and McEliece [25] by averaging over all possible binary representations of the RS code. We use the same technique here to analyze the performance of the sphere decoders, where the average binary weight enumerator of the RS code is used as the weight spectrum G_b of the binary linear code.

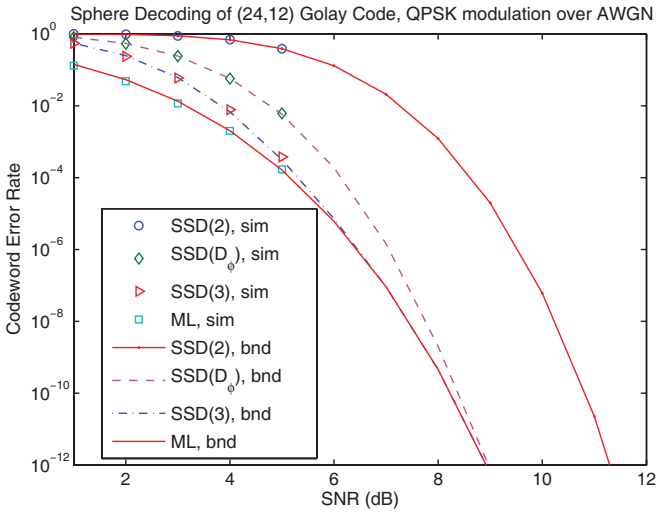

Fig. 3. Bounds on the performance of soft-decision sphere decoding of the $(24, 12)$ Golay code when QPSK modulated over an AWGN channel.

E. Numerical Results

In Fig. 3, we show how the bounds derived for M-ary modulated spherical codes are tight. The simulation curves and the analytical bounds (Theorem 5) will be labeled by 'sim' and 'bnd' respectively. A codeword in the $(24, 12)$ Golay code is mapped into 12 QPSK symbols and transmitted over an AWGN channel. As observed, the simulated performance of the ML decoder and the SD sphere decoder [14] are tightly bounded by the bounds given in this section. The critical decoding radius in the 2×12 dimensional space is $D_\phi = 2.667$.

In Fig. 4, the performance of SD sphere decoding of the binary image of the $(15, 11)$ RS code, BPSK modulated over an AWGN channel, is investigated. The ML performance is simulated by means of the MAP decoder, and it is observed that the averaged ML bound is tight [25]. We simulated the performance of SD sphere decoding when the decoding radius was 3 and 3.5 respectively. The analytical bounds of Theorem 5 almost overlapped with the simulations. The critical decoding radius is $D_\phi = 4.588$. A decoder with an Euclidean decoding radius of 5 has a near ML performance at an SNR of 5 dB. For reference purposes, we plot the performance of the hard-decision Berlekamp-Massey (BM) decoder. The bounds of Theorem 5 are compared to the conventional bound of Lemma 2 when the tight TSB is used as $\bar{P}(\mathcal{E}_{ML})$. It is observed that Theorem 5 offers a much tighter bound especially at low SNRs, where the bounds of Lemma 2 diverge and exceed one. The gap between the bounds is obvious at large decoding radii.

III. Sphere Decoding of Lattices

In this section, we consider the case of soft decision sphere decoding of a general lattice or code $\mathcal{C}$. In contrast to the case of section II the code is not constrained to be a linear code and the transmitted signal points (codewords) do not necessarily have the same energy. Define $G_w(i)$ to be the number of mapped codewords with an Euclidean distance δ_w from the ith codeword. Given that $\boldsymbol{c}_i$ is transmitted, let the

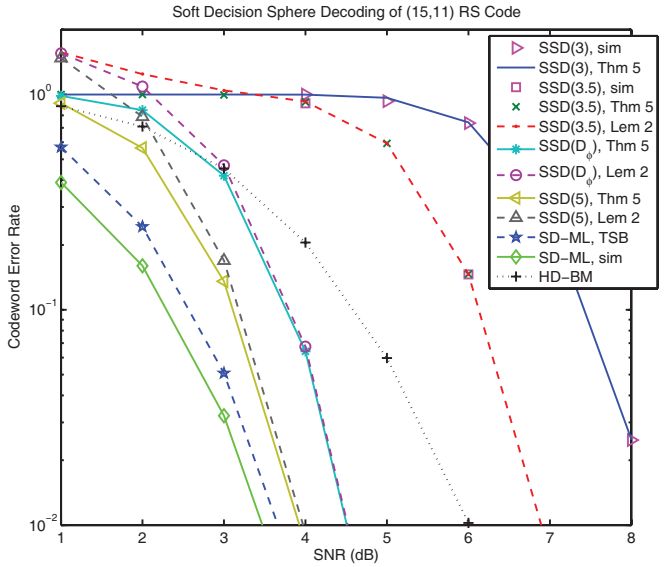

Fig. 4. Bounds on the performance of SSD of a binary image of the $(15, 11)$ Reed Solomon code BPSK modulated on an AWGN channel.

error probability of SSD(D) be upper bounded by $P_i(\mathcal{E}_D)$. By taking the expectation over all codewords,

$$P(\mathcal{E}_D) \leq \frac{1}{|\mathcal{C}|} \sum_{\boldsymbol{c}_i \in \mathcal{C}} P_i(\mathcal{E}_D). \tag{23}$$

Now, if we assume that $P_i(\mathcal{E}_D)$ is of the union bound form; $P_i(\mathcal{E}_D) = \sum_w G_w(i) P_i^{(w)}(\mathcal{E}_D)$, where $P_i^{(w)}(\mathcal{E}_D)$ is the probability of a sphere decoder error due to incorrectly decoding a codeword at a distance δ_w when $\boldsymbol{c}_i$ is transmitted. The error probability of SSD(D) can thus be upper bounded by $P(\mathcal{E}_D) \leq \sum_{\delta_w > 0} \bar{G}_w P^{(w)}(\mathcal{E}_D)$, where $P^{(w)}(\mathcal{E}_D)$ is the probability that the sphere decoder erroneously decodes a codeword at a distance δ_w from the transmitted codeword and

$$\bar{G}_w = \frac{1}{|\mathcal{C}|} \sum_{\boldsymbol{c}_i \in \mathcal{C}} G_w(i), \tag{24}$$

is the average number of codewords which are at an Euclidean distance δ_w from another codeword. For an arbitrary finite code or lattice $\mathcal{C}$, using arguments from the previous sections, the error probability SSD(D) can be upper bounded by

$$P(\mathcal{E}_D) \leq \min_{D' \leq D} \left\{ P(\mathcal{E}_{ML}, \boldsymbol{z} \in \Omega_{D'}) + P(\boldsymbol{z} \notin \Omega_{D'}) \right\}, \tag{25}$$

where $P(\boldsymbol{z} \notin \Omega_D)$ is given by (4) and

$$P(\mathcal{E}_{ML}, \boldsymbol{z} \in \Omega_D) = \sum_{0 < \frac{\delta_w}{2} < D} \bar{G}_w \int_{\frac{\delta_w}{2}}^{D} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-z^2/2\sigma^2} \Gamma_r\left(\frac{n_d - 1}{2}, \frac{D^2 - z^2}{2\sigma^2}\right) \mathrm{d}z. \tag{26}$$

The Hughes upper bound on the ML error probability is [26]

$$P(\mathcal{E}_{ML}) \leq \min_D P(\Psi(D)),$$

where

$$\Psi(D) \triangleq P(\mathcal{E}_{ML}, \boldsymbol{z} \in \Omega_D) + P(\boldsymbol{z} \notin \Omega_D). \tag{27}$$

The radius D_o that minimizes this error probability is the root of the equation [27]

$$\sum_{0 < \frac{\delta_w}{2} < D} \bar{G}_w \int_0^{\theta_{w,D}} \sin(\theta)^{n_d - 2} \mathrm{d}\theta = \frac{\sqrt{\pi}\Gamma\left(\frac{n_d - 1}{2}\right)}{\Gamma\left(\frac{n_d}{2}\right)}, \tag{28}$$

where $\theta_{w,d} = \cos^{-1}(\delta_w / 2D)$. From (25), the upper bound on the sphere decoding error probability is given by

$$P(\mathcal{E}_D) \leq \begin{cases} \Psi(D), & D < D_o \\ \Psi(D_o), & D \geq D_o \end{cases}.$$

Furthermore, the optimum radius D_o does not depend on the channel and can be the radius of choice for near maximum likelihood decoding. It is to be noted that the Hughes bound on ML decoding is not tighter than the Poltyrev tangential sphere bound [28].

For the case of M-PSK modulation of a linear code, the constellation may not result in a Hamming space if $M > 4$. In such a case the ensemble average weight enumerator $\bar{G}_w$ can be used with the bounds of Sec. II to analyze the performance. (The same technique can also be used with the results in next sections.)

Example 6: Assume an $(15, 3)$ RS code over F_{16} and assume a one-to-one mapping from the symbols of F_{16} to the points of an 16-QAM modulation [18], whose average energy per symbol is 10. The ensemble weight enumerator $\bar{G}_w$ was numerically computed to evaluate the bounds. The radius that minimizes the bound on the ML error probability is $D_o = 12.9$. In Fig. 5, we confirm that the bounds on the sphere decoder error probability agree with the simulations for the case of $D = 10$. We also compare the simulated performance of ML error probability $P(\mathcal{E}_{ML}, \boldsymbol{z} \in \Omega_D)$ to that of the analytic performance in both cases. At low SNRs this probability is low as the probability of the received word falling inside the sphere is relatively low. As more received words fall inside the sphere, the ML error probability increases as the SNR increases. At a certain SNR, the probability of the ML error starts decreasing due to the improved reliability of the received word.

IV. Performance of Sphere Decoding on Binary Symmetric Channels

In this section, an upper bound on the performance of the hard-decision sphere decoder, when the code is transmitted over the BSC, is derived. Transmitting a binary codeword over a binary input AWGN channel followed by hard decisions is equivalent to transmitting it on a BSC with a cross over probability $p = Q(\sqrt{2R\gamma})$ where γ is the bit signal to noise ratio. In case of M-PSK signaling with gray encoding, $p \approx \frac{p_c}{\log_2(M)}$ where $p_c = 2Q\left(\sqrt{2k\gamma} \sin\frac{\pi}{M}\right)$ [18].

Let $\boldsymbol{y}$ be the received word when the codeword $\boldsymbol{c}$ is transmitted over an BSC channel. The HD sphere decoder with radius $m - 1$, HSD$(m - 1)$, finds the codeword $\hat{\boldsymbol{c}}$, if it exists, such that

$$\begin{aligned} \hat{\boldsymbol{c}} = \arg\min_{\boldsymbol{v} \in \mathcal{C}} \quad & d(\boldsymbol{y}, \boldsymbol{v}) \\ \text{subject to} \quad & d(\boldsymbol{y}, \boldsymbol{v}) < m, \end{aligned} \tag{29}$$

where $d(\boldsymbol{y}, \boldsymbol{v})$ is the Hamming distance between $\boldsymbol{y}$ and $\boldsymbol{v}$. Let $\zeta = d(\boldsymbol{y}, \boldsymbol{c})$, then, from the linearity of the code, the

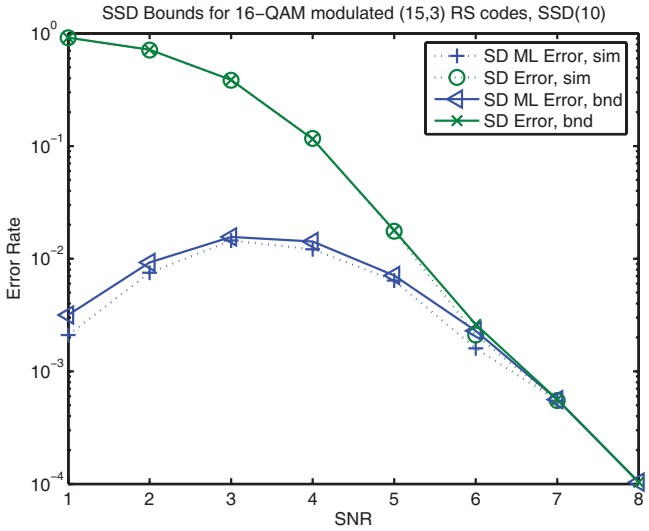

Fig. 5. The $(15, 3)$ RS code is 16-QAM modulated and transmitted over an AWGN channel. The sphere decoder is a soft decision sphere decoder with an Euclidean radius 10.The bounds are compared to simulations for a sphere decoding ML error, 'SD ML Error', and the error plus failure probability, 'SD Error'.

probability that the received word is outside a Hamming sphere (ball) of radius $m-1$ centered around the transmitted codeword is

$$P(\zeta \geq m) = \sum_{t=m}^{n} \binom{n}{t} p^t (1-p)^{n-t}. \tag{30}$$

Poltyrev [2] derived a tight bound on the performance of the HD-ML decoder based on,

$$P(\mathcal{E}_{ML}) \leq \min_{m} \left\{ P(\mathcal{E}_{ML}, \zeta < m) + P(\zeta \geq m) \right\}. \tag{31}$$

The minimum of the above equation is at m_o where m_o is the smallest integer m such that [2]

$$\sum_{b=1}^{2m} G_b \sum_{r=\lceil \frac{w}{2} \rceil}^{m} \binom{b}{r} \binom{n-b}{m-r} \geq \binom{n}{m}. \tag{32}$$

We now turn our attention to the HD sphere decoder with an arbitrary decoding radius. Let $P(\Sigma_m)$ be the error plus failure probability of the hard decision sphere decoder, HSD$(m-1)$, then $P(\Sigma_m)$ could be written as

$$\begin{aligned} P(\Sigma_m) &= P(\Sigma_m, \zeta < m) + P(\Sigma_m | \zeta \geq m) P(\zeta \geq m) \\ &= P(\mathcal{E}_{ML}, \zeta < m) + P(\zeta \geq m), \end{aligned} \tag{33}$$

where we used the fact that $P(\Sigma_m | \zeta \geq m) = 1$ and the observation that for $\zeta < m$, the conditional error probability of the HSD$(m-1)$ and the HD-ML decoders are the same. The last term in the above equation is a lower bound on the failure probability of the HSD$(m-1)$ decoder. Similar to the soft decision case, we have the following lemma,

Lemma 7: A lower bound on the performance of a hard decision sphere decoder, HSD$(m-1)$, over a BSC with parameter p is $P(\Sigma_m) \geq \sum_{t=m}^{n} \binom{n}{t} p^t (1-p)^{n-t}$.

To develop a tight upper bound on $P(\Sigma_m)$, we consider two cases:

Case I: The decoding radius $m \geq m_o$. Equation (33) could be written as

$$\begin{aligned} &P(\Sigma_m | m \geq m_o) = \\ &P(\mathcal{E}_{ML}, \zeta < m_o) + P(\mathcal{E}_{ML}, m_o \leq \zeta < m) + P(\zeta \geq m). \end{aligned}$$

It follows that

$$P(\Sigma_m | m \geq m_o) \leq P(\mathcal{E}_{ML}, \zeta < m_o) + P(\zeta \geq m_o). \tag{34}$$

We observe that the upper bound reduces to that of the HD-ML case (31). By recalling that the minimum of (31) is achieved at m_o, the bound of (33) is looser than (34) when $m > m_o$. The intuition behind this is that the performance of a sphere decoder with a decoding radius $m_o - 1$ or greater approaches that of the ML decoder.

Case II: The decoding radius $m < m_o$. Noticing that the sphere $\{\zeta < m\} \subset \{\zeta < m_o\}$, $P(\Sigma_m | m < m_o)$ is indeed given by (33).

Thus, we have proved the following theorem,

Theorem 8: The performance of a hard-decision sphere decoder with a decoding radius $m-1$ when used for decoding a linear code with a weight spectrum G_b over an BSC channel with a cross-over probability p is upper bounded by

$$\begin{aligned} &P(\Sigma_m) \leq \\ &\quad \begin{cases} P(\mathcal{E}_{ML}, \zeta < m_o) + P(\zeta \geq m_o), & m \geq m_o \\ P(\mathcal{E}_{ML}, \zeta < m) + P(\zeta \geq m), & m < m_o, \end{cases} \end{aligned} \tag{35}$$

where m_o is radius that minimizes (31) and is the solution of (32). $P(\zeta \geq m)$ is given by (30) and the joint probability of an HD-ML error and $d(\boldsymbol{y}, \boldsymbol{c}) < m$ is upper bounded by the union bound, $P(\mathcal{E}_{ML}, \zeta < m) \leq \sum_{b=1}^{2(m-1)} G_b \sum_{r=\lceil \frac{w}{2} \rceil}^{m-1} \left[\binom{b}{r} p^r (1-p)^{b-r} \sum_{s=0}^{m-r-1} \binom{n-b}{s} p^s (1-p)^{n-b-s} \right]$. ▽

A. Numerical Examples

In this subsection, the bounds developed for SD and HD sphere decoding are evaluated and compared with the performance of the corresponding sphere decoders, [14] and [15] respectively.

In Fig. 6, we compare the analytical bounds to simulations of sphere decoding of an $(15, 7)$ BCH code BPSK modulated and transmitted over an AWGN channel. The minimum distance of the BCH code is 5. The critical Euclidian decoding radius of the soft decision decoder is $D_\phi = 3.17$ while the critical Hamming decoding radius of the hard decision decoder is $m_o = 3$. We observe that the simulated performance is tightly upper bounded by the analytical bounds of theorems 4 and 8 for soft and hard decision sphere decoding respectively. The larger the decoding radius the nearer the performance is to maximum likelihood decoding.

V. Performance of Sphere Decoding on Q-ary Symmetric Channels

Now consider an (n, k, d) RS code and a hard-decision sphere decoder which can correct τ symbol errors, where the symbols are in F_q. The Berlekamp-Massey algorithm is a well known polynomial time algorithm that can correctly decode words which are at a (symbol) Hamming distance of

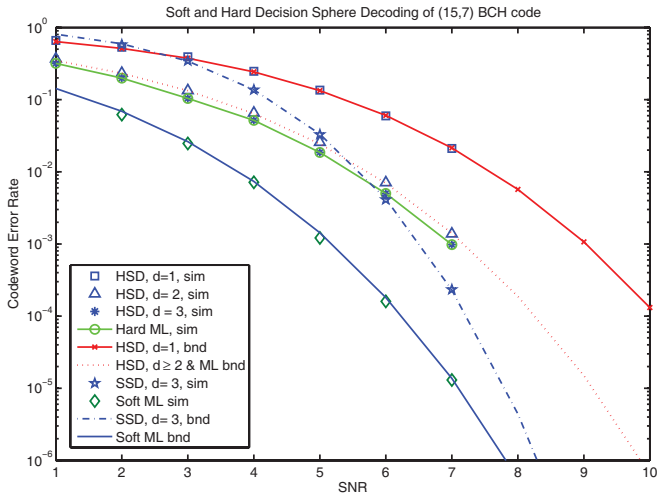

Fig. 6. Bounds on the codeword error rate of hard decision sphere decoding (HSD) and soft decision sphere decoding (SSD) of the $(15, 7)$ BCH code BPSK modulated over an AWGN channel. The simulations (labeled by 'sim') are tightly upper bounded by the analytic bounds (labeled by 'bnd') .

$\tau_{BM} = \lfloor \frac{n-k}{2} \rfloor$ from the transmitted codeword. The error probability of bounded distance decoding of RS codes is well studied (cf. [29]). Recently, Guruswami and Sudan [16] developed a list decoding algorithm that can correct upto $\tau_{GS} = \lceil n - \sqrt{nk} - 1 \rceil$ symbol errors. To analyze this case, we first derive a bound on the performance of the corresponding ML decoder.

A. Bound on the Maximum Likelihood decoding of linear block codes on q-ary symmetric channels.

We will assume an (n, k, d) linear code over F_q transmitted over a q-ary symmetric channel. The probability that a symbol is correctly received will be denoted by s, while the probability that it is received as another symbol will be $p = (1-s)/(q-1)$. We assume $q = 2^m$, the channel alphabet size is 2^b, $b \leq m$ and each q-ary symbol is mapped to m/b channel symbols. If p_c is the probability that a channel symbol is incorrectly decoded, then $s = (1 - p_c)^{m/b}$ [18].

Let ζ be the Hamming distance between the transmitted codeword and the received q-ary word. Then, similar to the binary case, the ML error probability can be upper bounded as follows,

$$P(\mathcal{E}_{ML}) \leq \min_{m} \{P(\mathcal{E}_{ML}, \zeta < m) + P(\zeta \geq m)\}. \quad (36)$$

Assuming that the code is linear, the probability that the received q-ary word lies outside a Hamming sphere (ball) of radius $m-1$ centered around the transmitted word is

$$P(\zeta \geq m) = \sum_{\alpha=m}^{n} \binom{n}{\alpha} (1-s)^{\alpha} s^{n-\alpha}. \quad (37)$$

The above equation also provides a lower bound on the performance of the sphere decoder HSD$(m-1)$.

The first term in (36) is upper bounded by the following lemma.

Lemma 9: For an (n, k, d) linear code over F_q, with a weight enumerator G_w, transmitted over a q-ary symmetric channel with parameters s and p,

$$P(\mathcal{E}_{ML}, \zeta < m) \leq \sum_{w=d}^{2(m-1)} G_w \sum_{\alpha=0}^{m-1} \sum_{\eta=\lceil \frac{w-\alpha}{2} \rceil}^{w-\alpha} \left(\frac{w!}{\eta!\alpha!(w-\eta-\alpha)!} p^{\eta} (1-p-s)^{\alpha} s^{w-\eta-\alpha} \sum_{\beta=0}^{m-1-\eta-\alpha} \binom{n-w}{\beta} (1-s)^{\beta} s^{n-w-\beta} \right). \quad (38)$$

Proof: We will assume that the all-zero codeword is transmitted. Now consider a codeword $\boldsymbol{c}$ with Hamming weight w and assume the received word $\boldsymbol{r}$ has a Hamming weight $m' - 1$. Consider the w nonzero symbols in $\boldsymbol{c}$ and the corresponding coordinates in $\boldsymbol{r}$. Let $\boldsymbol{r}$ and $\boldsymbol{c}$ have the same symbols in η of these coordinates. Let α of these w coordinates in $\boldsymbol{r}$ be neither zero nor match those in $\boldsymbol{c}$, and $w - \eta - \alpha$ of the remaining coordinates be zero. Since the Hamming weight of $\boldsymbol{r}$ is $m'-1$, there must be $m'-1-\eta-\alpha$ non-zero symbols in the remaining $n-w$ coordinates and the remaining symbols will be zero. The probability of receiving such a word is $\frac{w!}{\eta!\alpha!(w-\eta-\alpha)!} p^{\eta} (1-p-s)^{\alpha} s^{w-\eta-\alpha} \binom{n-w}{m'-1-\eta-\alpha} (1-s)^{m'-1-\eta-\alpha} s^{n-w-(m'-1-\eta-\alpha)}$. In such a case, the Hamming distance between $\boldsymbol{r}$ and $\boldsymbol{c}$ is $w + m' - 1 - 2\eta - \alpha$. An ML error results if this is less than the weight of $\boldsymbol{r}$, i.e., if $\eta \geq \lceil \frac{w-\alpha}{2} \rceil$. By summing over all possible combinations of η and α and applying the union bound for all codewords that can be within a Hamming distance m' from $\boldsymbol{r}$, the error probability is upper bounded by

$$\sum_{w=d}^{2(m'-1)} G_w \sum_{\alpha=0}^{m'-1} \sum_{\eta=\lceil \frac{w-\alpha}{2} \rceil}^{w-\alpha} \left(\frac{w!}{\eta!\alpha!(w-\eta-\alpha)!} p^{\eta} (1-p-s)^{\alpha} s^{w-\eta-\alpha} \binom{n-w}{m'-1-\eta-\alpha} (1-s)^{\beta} s^{n-w-(m'-1-\eta-\alpha)} \right).$$

Applying the union bound for all received words with Hamming weights less than m, $m' \leq m$, the result follows. ■

We are now ready to prove the following theorem,

Theorem 10: The ML error probability of an (n, k, d) q-ary linear code on a q-ary symmetric channel is upper bounded by

$$P(\mathcal{E}_{ML}) \leq \sum_{w=d}^{2(m_o-1)} G_w \sum_{\alpha=0}^{m_o-1} \sum_{\eta=\lceil \frac{w-\alpha}{2} \rceil}^{w-\alpha} \left(\frac{w!}{\eta!\alpha!(w-\eta-\alpha)!} p^{\eta} (1-p-s)^{\alpha} s^{w-\eta-\alpha} \sum_{\beta=0}^{m_o-1-\eta-\alpha} \binom{n-w}{\beta} (1-s)^{\beta} s^{n-w-\beta} \right) + \sum_{\alpha=m_o}^{n} \binom{n}{\alpha} (1-s)^{\alpha} s^{n-\alpha}, \quad (39)$$

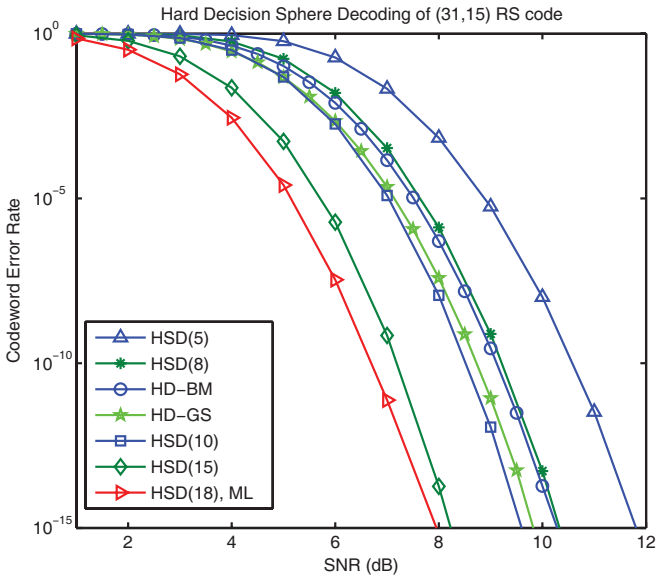

Fig. 7. Bounds on the performance of hard decision sphere decoding of the $(31, 15)$ RS code BPSK on an AWGN channel.

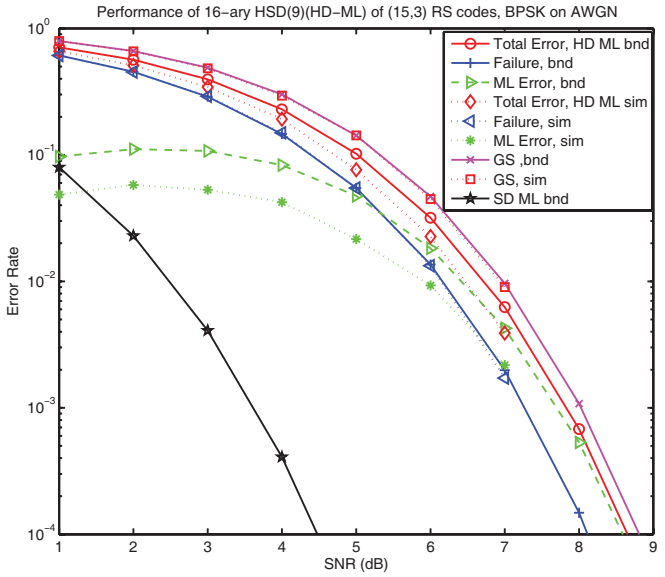

Fig. 8. The $(15, 3)$ RS code is BPSK modulated and transmitted over an AWGN channel. For the 16-ary hard-decision decoder, the channel is an QSC. The bounds are compared to simulations for a sphere decoding ML error, sphere decoding failure, and their sum (Total Error, HD ML). The optimum radius for the ML bound is 9. The GS radius is 8.

where m_o is the smallest integer m such that

$$\sum_{w=d}^{2m} G_w \sum_{\alpha=0}^{m} \left(\frac{q-2}{q-1}\right)^{\alpha} \sum_{\eta=\lceil \frac{w-\alpha}{2} \rceil}^{w-\alpha} \left(\frac{1}{q-1}\right)^{\eta} \frac{w!}{\eta!\alpha!(w-\eta-\alpha)!} \binom{n-w}{m-\eta-\alpha} \geq \binom{n}{m}. \quad (40)$$

▽

Proof: The upper bound follows by substituting (38) and (37) in (36). Observe that the first term in (39) is increasing in m while the second is decreasing in m. Optimizing over the radius m, the minimum is achieved at the first integer m such that $\sum_{w=d}^{2(m)} G_w \sum_{\alpha=0}^{m} \sum_{\eta=\lceil \frac{w-\alpha}{2} \rceil}^{w-\alpha} \left(\frac{w!}{\eta!\alpha!(w-\eta-\alpha)!} \; p^{\eta} (1-p-s)^{\alpha} s^{w-\eta-\alpha} \; \binom{n-w}{m-\eta-\alpha} (1-s)^{m-\eta-\alpha} s^{n-w-m+\eta+\alpha} \right) \geq \binom{n}{m} (1-s)^{m} s^{n-m}$. This reduces to the condition of (40). ■

It is worth noting that the optimum radius m_o which minimizes the bound on the ML error probability only depends on the weight enumerator of the code and the size of its finite field. Since the optimum radius does not depend on the SNR, it is valid for q-ary symmetric channels at any SNR. Similar to the binary case [2], we establish below a connection between m_o and the covering radius of the code.

Lemma 11: The covering radius of a linear code on F_q is lower bounded by $m_o - 1$, where m_o is given by Theorem 10.

Proof: Define $L(m)$ to be the left hand side term in (40) and $\boldsymbol{c}_o$ to be the all zero codeword. Similar to the proof of Lemma 9, one can show that

$$(q-1)^m L(m) = |\{\boldsymbol{r} \in F_q^n : d(\boldsymbol{r}, \boldsymbol{c}_o) = m \;\; \& \;\; d(\boldsymbol{r}, \boldsymbol{c}_i) \leq m; \quad \boldsymbol{c}_i \in \mathcal{C} \setminus \boldsymbol{c}_o\}|.$$

Also, $(q-1)^m \binom{n}{m} = |\{\boldsymbol{r} \in F_q^n : d(\boldsymbol{r}, \boldsymbol{c}_o) = m\}|$. Since

$$(q-1)^{m_o-1} L(m_o - 1) < (q-1)^{m_o-1} \binom{n}{m_o - 1},$$

it follows that there exit words $\boldsymbol{r} \in F_q^n$ such that $\min_{\boldsymbol{c} \in \mathcal{C}} d(\boldsymbol{r}, \boldsymbol{c}) = m_o - 1$ and this minimum is achieved when $\boldsymbol{c}$ is the all zero codeword $\boldsymbol{c}_o$. By recalling that the covering radius is [30]

$$R_c = \max_{\boldsymbol{r} \in F_q^n} \min_{\boldsymbol{c} \in \mathcal{C}} d(\boldsymbol{r}, \boldsymbol{c}),$$

it follows that $R_c \geq m_o - 1$. ■

B. Hard Decision Sphere decoding of linear block codes on q-ary symmetric channels.

Here, we consider the case when the decoder is a q-ary hard decision sphere decoder. As for the binary case, the HSD$(m-1)$ can correctly decode a codeword if the number of q-ary symbol errors is $m-1$ or less. Thus Theorem 8 will give the bound on the error plus failure probability of the sphere decoder. However, in this case, $P(\zeta \geq m)$, $P(\mathcal{E}_{ML}, \zeta < m)$ and m_o are given by (37), (38) and (40) respectively.

C. Numerical Examples

In Fig. 7, we show bounds on the performance of HD decoding of the near half rate $(31, 15)$ RS code over F_{32} when its binary image is transmitted over an AWGN channel followed by hard-decisions. The optimum binary decoding radius is 18. Thus the closer the decoding radius is to 18, the better the performance of the sphere decoder. The HD-ML decoder has more than 2 dB coding gain over the Berlekamp Massey (BM) decoder, which can correct 8 symbol errors. It is observed that the average performance of an HD sphere decoder, with a (binary Hamming) radius 8, closely upper bounds that of the HD-BM decoder that can correct 8 symbol errors. The HD-GS decoder can correct one more symbol error than the BM decoder. The performance of the GS algorithm is analyzed by modeling it as 16-ary HD sphere decoder of radius 9. Consequently, one can observe that a hard-decision sphere

decoder with a binary decoding radius of 10 outperforms the symbol based GS decoder.

In Fig. 8, the binary image of the $(15, 3)$ RS code is BPSK modulated over an AWGN channel. For 16-ary hard decisions, the channel is modeled as an QSC. The performance bound of the hard ML (HD ML) decoder is shown (Theorem 10) and is the same as an HSD of radius 9. The bounds of (37) and (38) are also shown and labeled as $F(9)$ and $E(9)$ respectively. As seen, the three bounds ('bnd') are in close agreement with the simulation ('sim'), for such a hypothetical sphere decoder. The error probability of the GS decoder with radius 8 is simulated and agrees with the bound of Theorem 8. For reference purposes, we show the average error probability of the soft decision bit level ML (SD ML) decoder (cf [25]) which has about 4 dB gain over the symbol HD ML decoder.

VI. A Note on Complexity

In Fig. 9, the empirical complexity exponents of SSD of the $(24, 12)$ Golay code BPSK modulated over an AWGN channel are shown. It is clear that for a larger decoding radius there is a price paid in terms of the complexity. We also show the complexity of the SSD whose radius changes such that with a probability of 0.9 the transmitted word is inside the sphere centered around the received one. At a slight increase in average complexity one can achieve ML decoding, by gradually increasing the decoding radius until a word is found. The corresponding complexity is shown as '$r^2 0.90 +$ cumulative'. The variation of the radius versus the SNR is shown in Fig. 10.

VII. Conclusions

Bounds on the error plus failure probability of hard-decision and soft-decision sphere decoding of block codes were derived. By comparing with the simulations of the corresponding decoders, we demonstrate that our bounds are tight. The ML performance of codes on q-ary symmetric channels is analyzed. The performance of sphere decoding of Reed Solomon codes and their binary images was analyzed. Moreover, the bounds are extremely useful in predicting the performance of the sphere decoders at the tail of error probability when simulations are prohibitive. The bounds allows one to pick the radius of the sphere decoder that best fits the performance, throughput and complexity requirements of the system.

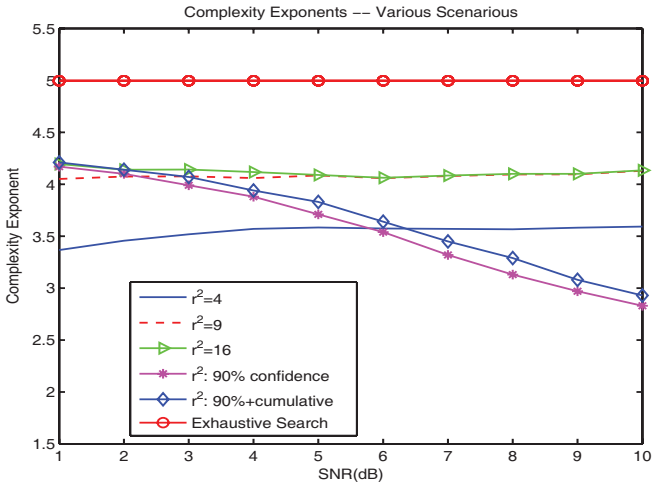

Fig. 9. Complexity exponent for SSD of the $(24, 12)$ Golay code.

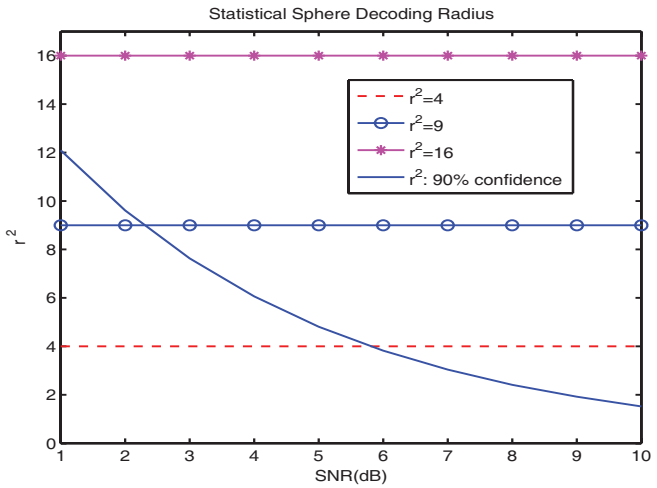

Fig. 10. Statistical Decoding Radius vs Fixed Decoding Radius for the $(24, 12)$ Golay code.

Acknowledgment

We thank the reviewers and the associate editor for their comments that significantly improved the quality of this paper.

References

[1] Berlekamp, R. McEliece, and H. van Tilborg, "On the inherent intractability of certain coding problems," *IEEE Trans. Inform. Theory*, vol. 24, pp. 384-386, May 1978.

[2] G. Poltyrev, "Bounds on the decoding error probability of binary linear codes via their spectra," *IEEE Trans. Inform. Theory*, vol. 40, no. 4, pp. 1284-1292, July 1994.

[3] E. Berlekamp, "The technology of error-correcting codes," in *Proc. IEEE*, vol. 68, no. 8, pp. 564-593, May 1980.

[4] S. Aji, H. Jin, A. Khandekar, D. J. Mackay, and R. J. McEliece, "BSC thresholds for code ensembles based on "typical pairs" decoding," in *Proc. IMA Workshop Codes Graphs*, Aug. 1999, pp. 195-210.

[5] D. Divsalar and E. Biglieri, "Upper bounds to error probabilities of coded systems over AWGN and fading channels," in *Proc. 2000 IEEE Global Telecommun. Conf. (GLOBECOM'00), San Francisco, CA*, Nov. 2000, pp. 1605-1610.

[6] I. Sason, S. Shamai, and D. Divsalar, "Tight exponential upper bounds on the ML decoding error probability of block codes over fully interleaved fading channels," *IEEE Trans. Commun.*, vol. 51, no. 8, pp. 1296-1305, Aug. 2003.

[7] I. Sason and S. Shamai, "Performance analysis of linear codes under maximum-likelihood decoding: a tutorial," *Foundations Trends Commun. Inform. Theory*, vol. 3, July 2006.

[8] U. Fincke and M. Pohst, "Improved methods for calculating vectors of short length in a lattice, including a complexity analysis," *Mathematics Computation*, vol. 44, pp. 463-471, 1985.

[9] C. Schnorr and M. Euchner, "Lattice basis reduction: improved practical algorithms and solving subset sum problems," *Math. Programming*, vol. 66, pp. 181-191, 1994.

[10] E. Agrell, A. Vardy, and K. Zeger, "Closest point search in lattices," *IEEE Trans. Inform. Theory*, vol. 48, no. 8, pp. 2201-2214, Aug. 2002.

[11] E. Viterbo and J. Boutros, "A universal lattice decoder for fading channels," *IEEE Trans. Inform. Theory*, vol. 45, p. 1639.

[12] M. O. Damen, A. Chkeif, and J. Belfiore, "Lattice code decoder for space-time codes," *IEEE Commun. Lett.*, pp. 161-163, May 2000.

[13] M. O. Damen, H. E. Gamal, and G. Caire, "On maximum-likelihood detection and the search for the closest lattice point," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2389-2402, 2003.

[14] H. Vikalo and B. Hassibi, "On joint detection and decoding of linear block codes on Gaussian vector channels," *IEEE Trans. Signal Processing*, Sept. 2006.

[15] ——, "Statistical approach to ML decoding of linear block codes on symmetric channels," in *Proc. IEEE International Symp. Inform. Theory (ISIT)*, 2004.
[16] V. Guruswami and M. Sudan, "Improved decoding of Reed-Solomon codes and algebraic geometry codes," *IEEE Trans. Inform. Theory*, vol. 45, no. 6, pp. 1757-1767, Sept. 1999.
[17] H. Vikalo, B. Hassibi, and U. Mitra, "Sphere-constrained ML detection for frequency-selective channels," *IEEE Trans. Commun.*, no. 7, pp. 1179-1183, 2006.
[18] J. G. Proakis, *Digital Communications*, 4th ed. New York: McGraw-Hill, 2001.
[19] Hochwald and S. ten Brink, "Achieving near-capacity on a multiple-antenna channel," *IEEE Trans. Commun.*, vol. 53, pp. 389-399, Mar. 2003.
[20] G. D. Forney, "Geometrically uniform codes," *IEEE Trans. Inform. Theory*, vol. 37, no. 5, pp. 1241-1260, Sept. 1991.
[21] E. W. Weisstein, *Mathworld–A Wolfram Web Resource*. [Online]. Available: http://mathworld.wolfram.com.
[22] C. E. Shannon, "Probability of error for optimal codes in a Gaussian channel," *Bell Syst. Tech. J.*, vol. 38, pp. 611-656, 1959.
[23] H. Herzberg and G. Poltyrev, "The error probability of M-ary PSK block coded modulation schemes," *IEEE Trans. Commun.*, vol. 44, no. 4, pp. 427-433, Apr. 1996.
[24] C. Retter, "The average binary weight enumerator for a class of generalized Reed-Solomon codes," *IEEE Trans. Inform. Theory*, vol. 37, no. 2, pp. 346-349, Mar. 1991.
[25] M. El-Khamy and R. J. McEliece, "Bounds on the average binary minimum distance and the maximum-likelihood performance of Reed Solomon codes," in *42nd Allerton Conf. Commun., Control Computing*, 2004.
[26] B. Hughes, "On the error probability of signals in additive white Gaussian noise," *IEEE Trans. Inform. Theory*, pp. 151-155, Jan. 1991.
[27] H. Herzberg and G. Poltyrev, "Techniques of bounding the probability of decoding error for block coded modulation structures," *IEEE Trans. Inform. Theory*, pp. 903-911, May 1994.
[28] D. Divsalar, "A simple tight bound on error probability of block codes with application to turbo codes," TMO Progress Report, NASA, JPL, Tech. Rep. 42-139, 1999.
[29] R. J. McEliece and L. Swanson, "On the decoder error probability of Reed-Solomon codes," *IEEE Trans. Inform. Theory*, vol. 32, no. 5, pp. 701-703, Sept. 1986.
[30] F. J. MacWilliams and N. J. Sloane, *The Theory of Error Correcting Codes*. Amsterdam: North Holland, 1977.

Mostafa El-Khamy received the Ph.D. degree in electrical engineering in 2006 from the California Institute of Technology (Caltech), Pasadena, USA where he also received an M.S. degree in electrical engineering in 2003. He received the B.S. degree (with honor) and an M.S. degree in electrical engineering (communications and electronics section) from Alexandria University, Egypt in 1999 and 2001 respectively. He has been a senior systems engineer in Corporate Research and Development at Qualcomm, San Diego, USA from October 2006 to October 2008 where he has contributed to the design and development of next generation mobile wireless systems, with an emphasis on 3GPP WCDMA/HSPA+ evolution. He has also been a key member in the analysis and design of advanced solutions for the deployment of femto-cells and home access points. He is currently an assistant professor at Alexandria University, Egypt and a visiting lecturer at the German University in Cairo, Egypt.

He received the Atwood Fellowship from the California Institute of Technology in 2002. He is also the recipient of the Young Scientist Award from the URSI General Assembly, the Netherlands, in 2002. His research interests include wireless communications, information theory, coding theory and signal processing.

Haris Vikalo received his B.S. degree from the University of Zagreb, Croatia, in 1995, the M.S. degree from Lehigh University in 1997, and the Ph.D. degree from Stanford University in 2003, all in electrical engineering. He held a short-term appointment at Bell Laboratories, Murray Hill, NJ, in the summer of 1999. From January 2003 to July 2003 he was a Postdoctoral Researcher, and from July 2003 to August 2007 he was an Associate Scientist at the California Institute of Technology. Since September 2007, he has been with the Department of Electrical and Computer Engineering, the University of Texas at Austin, where he is currently an Assistant Professor. His research interests include genomic signals and systems, stochastic signal processing, wireless communications, and algorithm complexity.

Babak Hassibi was born in Tehran, Iran, in 1967. He received the B.S. degree from the University of Tehran in 1989, and the M.S. and Ph.D. degrees from Stanford University in 1993 and 1996, respectively, all in electrical engineering.

From October 1996 to October 1998 he was a research associate at the Information Systems Laboratory, Stanford University, and from November 1998 to December 2000 he was a Member of the Technical Staff in the Mathematical Sciences Research Center at Bell Laboratories, Murray Hill, NJ. Since January 2001 he has been with the California Institute of Technology, Pasadena, CA., where he is currently professor and executive officer of electrical engineering. He has also held short-term appointments at Ricoh California Research Center, the Indian Institute of Science, and Linkoping University, Sweden. His research interests include wireless communications, robust estimation and control, adaptive signal processing and linear algebra. He is the coauthor of the books *Indefinite Quadratic Estimation and Control: A Unified Approach to H^2 and H^∞ Theories* (New York: SIAM, 1999) and *Linear Estimation* (Englewood Cliffs, NJ: Prentice Hall, 2000). He is a recipient of an Alborz Foundation Fellowship, the 1999 O. Hugo Schuck best paper award of the American Automatic Control Council, the 2002 National Science Foundation Career Award, the 2002 Okawa Foundation Research Grant for Information and Telecommunications, the 2003 David and Lucille Packard Fellowship for Science and Engineering and the 2003 Presidential Early Career Award for Scientists and Engineers (PECASE), and was a participant in the 2004 National Academy of Engineering "Frontiers in Engineering" program.

He has been a Guest Editor for the IEEE TRANSACTIONS ON INFORMATION THEORY special issue on "space-time transmission, reception, coding and signal processing" was an Associate Editor for Communications of the IEEE TRANSACTIONS ON INFORMATION THEORY during 2004-2006, and is currently an Editor for the journal FOUNDATIONS AND TRENDS IN INFORMATION AND COMMUNICATION.

Robert J. McEliece has been on the faculty at the California Institute of Technology, Pasadena, since 1982, where he is now the Allen E. Puckett Professor and Professor of Electrical Engineering. From 1990 to 1999, he served as Executive Officer for Electrical Engineering at the California Institute of Technology. He has been a Consultant in the Communications Research Section of the Jet Propulsion Laboratory since 1978. His research interests include deep-space communication, communication networks, coding theory, and discrete mathematics.

Dr. McEliece is a member of the National Academy of Engineering. He was the recipient of the 2004 IEEE Information Theory Society Shannon Award.